

## University of Groningen

### DNA recombination in plants

Bagherieh Najjar, Mohammad Bagher

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

*Document Version*

Publisher's PDF, also known as Version of record

*Publication date:*

2004

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Bagherieh Najjar, M. B. (2004). *DNA recombination in plants: molecular and functional analysis of Arabidopsis RecQ genes*. [Thesis fully internal (DIV), University of Groningen]. s.n.

**Copyright**

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

**Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

# Chapter 2

## **Molecular characterisation of the *Arabidopsis RecQ*-like genes**

Mohammad B. Bagherieh-Najjar, Jacques Hille and Paul P. Dijkwel

## ABSTRACT

Members of the RecQ family of DNA helicases are found in different kingdoms of living organisms. Bacteria and yeasts contain a single *RecQ* gene and the human genome contains at least five. Studied RecQ proteins play a crucial role in aspects of DNA metabolism including replication, recombination, and repair. Here we report on the isolation and molecular characterization of *RecQ* genes from *Arabidopsis thaliana*, a model flowering plant. *Arabidopsis* contains seven *RecQ* genes (*AtRecQ*-like) with a relatively complex structure. A phylogenetic analysis revealed that *AtRecQ14A* and *AtRecQ14B* have evolved from a recent duplication but the others are more distantly related, suggesting that multiple *RecQ* genes were present in the last common ancestor of plants and animals. All seven *AtRecQ*-like genes were ectopically expressed in *Saccharomyces cerevisiae* to evaluate their ability to compensate for the absence of the sole budding yeast *RecQ* homologue, *SGS1*. *AtRecQ14A* almost fully suppressed the *sgs1* hypersensitivity to the DNA damaging drug methyl methanesulfonate (MMS) and hyperrecombination phenotypes. *AtRecQsim* partially suppressed the MMS hypersensitivity but enhanced the hyper-recombination phenotype. *AtRecQ13* did not have any effect, while the other four *AtRecQ*-like genes enhanced both MMS hypersensitivity and hyper-recombination phenotypes of *sgs1* cells. The data suggest that most *AtRecQ*-like genes acquired different functions from *SGS1* and that *RecQ14A* is the *SGS1* orthologue in flowering plants.

## INTRODUCTION

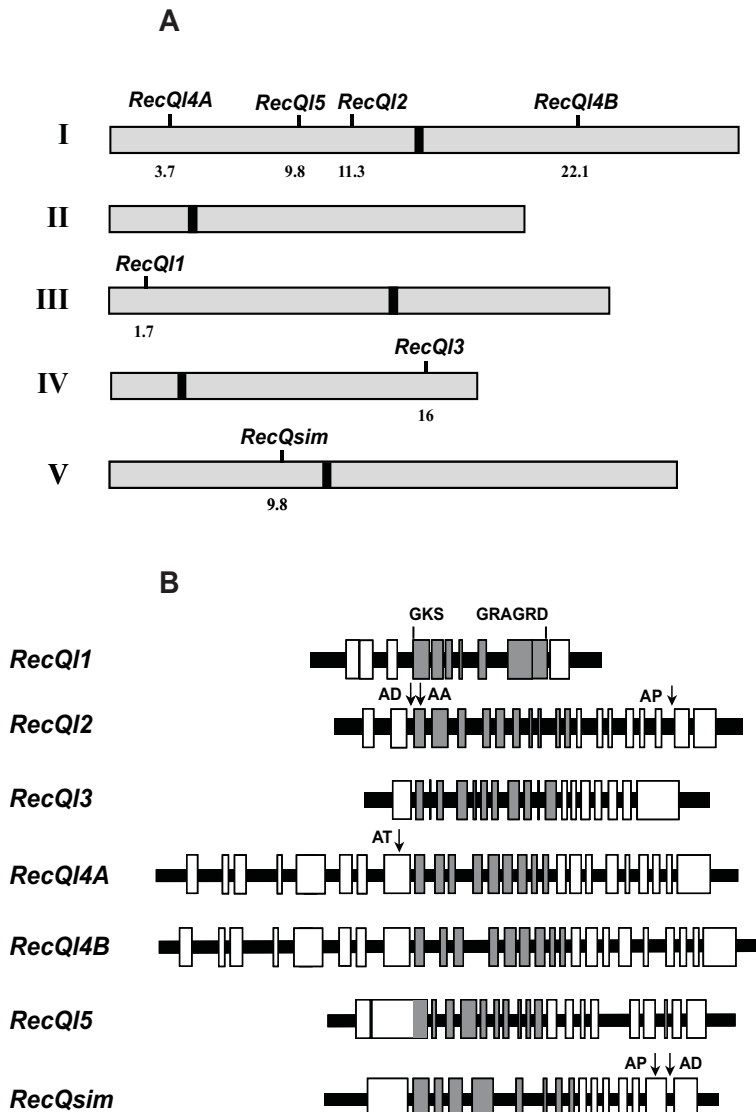
RecQ DNA helicases play vital roles in the maintenance of genome integrity in pro- and eukaryotes. The prototypical member of the family, *Escherichia coli* RecQ, is involved in the RecF pathway of homologous recombination and is a suppressor of illegitimate recombination (Nakayama et al., 1985; Hanada et al., 1997). *S. cerevisiae* contains a single *RecQ* gene, *SGS1*, which was initially isolated as a suppressor of the slow growth phenotype of *top3* mutants (Gangloff et al., 1994). Mutations in *SGS1* result in increased rate of recombination, reduced life span, impaired sporulation, enhanced sensitivity to the DNA synthesis inhibitor hydroxyurea (HU), and hypersensitivity to MMS (Gangloff et al., 1994; Watt et al., 1996; Sinclair and Guarente 1997; Mullen et al., 2000). The Human genome contains at least five *RecQ* family members, of which three are implicated in heritable diseases. Mutations in *BLM*, *WRN*, or *RecQ4* genes are associated with Bloom, Werner, or Rothmund-Thomson syndromes, respectively (Ellis et al., 1995; Yu et al., 1996; Kitao et al., 1999). In each case, cells derived from affected individuals show inherent genomic instability associated with several features of premature ageing and/or cancer predisposition (reviewed in Karow et al., 2000b; Bachrati and Hickson, 2003).

RecQ proteins share a highly conserved helicase domain, which comprises of approximately 350 amino acids. This central region includes seven motifs and a DEXHC-box in the second motif is a characteristic of the RecQ family (Kusano et al., 1999). Most

RecQ proteins contain a C-terminal extension that can be divided into two domains (Morozov et al., 1997). The RecQ-C terminal (RecQ-Ct) domain, which is unique to the RecQ family, is located proximal of the helicase domain. Mutations at several conserved residues within the RecQ-Ct domain of BLM reduced ATPase and helicase activities as well as single-stranded DNA-binding of the protein (Janscak et al., 2003). Next to the RecQ-Ct domain the RNaseD C-terminal (HRDC) domain can be found that also occurs in RNaseD proteins. It has been suggested that the HRDC domain of SGS1 and BLM modulates the helicase activity of the protein through single-stranded DNA binding, but it is not essential for the enzyme activity (Miyajima et al., 2000; Ui et al., 2001; Janscak et al., 2003; Bernstein and Keck, 2003). The diversity in the N-terminal region of the helicase domain is more than that in the C-terminal region. Some members, like *E. coli* RecQ and human RecQ5 do not have extended N-termini, while others such as WRN and BLM carry different domains at their N-terminal regions. The human WRN protein and its orthologues are unique in carrying an N-terminal exonuclease domain (Shen and Loeb, 2000). It is widely accepted that the N- and C-terminal regions play a key role in the functional specificities of RecQ proteins.

In addition to the structural homology, there is a certain degree of functional conservation among RecQ proteins. Where studied, they have been shown to be substrate specific ATP-dependent DNA helicases that translocate in the 3' to 5' direction (reviewed in Bachrati and Hickson, 2003). *In vitro* analysis revealed that RecQ proteins are capable of unwinding alternatively structured DNA molecules, including B-form, synthetic D-loops, Holliday junctions and G4 DNA (Sun et al., 1998; Sun et al., 1999; van Brabant et al., 2000; Karow et al., 2000a; Mohaghegh et al., 2001). Holliday junctions (HJ) are DNA intermediate structures that arise during DNA recombination and replication (Sharples et al., 1999). It has been suggested that HJ unwinding activity of BLM and WRN may result in reducing inappropriate DNA recombination intermediates *in vivo* (Constantinou et al., 2000; Karow et al., 2000a). G4 DNA is a stable four stranded structure, in which the repeating unit is a G-quartet stabilised by hydrogen bonds. This unusual DNA structure can be found mainly in two G-rich genomic domains, rDNA and telomeres, and is not a substrate for all helicases. SGS1 and BLM unwind G4 DNA with at least 15-fold preference relative to duplex substrates (Sun et al., 1998; Sun et al., 1999) and several-fold preference relative to Holliday junctions (Huber et al., 2002). The ability of resolving aberrant DNA structures, including G4-DNA and HJ may be an explanation of how a *RecQ* deficiency in human and yeast cells might lead to telomere deregulation associated with premature ageing and increased genomic instability. The functional conservation in RecQ proteins is further supported by data showing that ectopic expression of both *BLM* and *WRN* can suppress the enhanced illegitimate and homologous recombination phenotypes of the *sgs1* mutant. *BLM*, but not *WRN*, is able to restore the premature ageing phenotype of *sgs1* cells and decrease the growth rate of *sgs1 top3* cells to the *top3* level (Yamagata et al., 1998; Heo et al., 1999; Neff et al., 1999). In contrast to human *WRN*, *BLM* and *RecQ4* (Stein et al., 2002), the only studied *Neurospora crassa* *RecQ* member (*QDE3*) is involved in RNAi (Cogoni and Macino, 1999). The role of plant RecQ family members, however, is unknown.

Here we report on the isolation and molecular characterization of the seven Arabidopsis *RecQ*-like homologues followed by a complementation analysis in the yeast *sgs1* mutant. We demonstrate that Arabidopsis *RecQ*-like genes can play various roles in *sgs1* cells, which suggests their specialised functions in plant metabolism.



**Figure 1.** Molecular characterisation of the Arabidopsis *RecQ*-like genes. **(A)** Schematic representation of the map positions of the Arabidopsis *RecQ*-like genes. The chromosome numbers are indicated to the left and the black boxes represent the approximate positions of the centromeres. The numbers below each chromosome indicate the positions of the *AtRecQ*-like genes in mega base pairs. **(B)** Schematic genomic structure of the seven Arabidopsis *RecQ*-like genes. Exons are shown as boxes. Grey boxes indicate the exons that encode for the helicase domain. 'GKS' and 'GRAGRD' represent the conserved amino acids of the 1<sup>st</sup> and 7<sup>th</sup> motif of the helicase domain, respectively. Arrows show the position of the detected alternative mRNA processing events. AT, alternative transcription initiation site; AD, alternative donor site; AA, alternative acceptor site; AP, alternative polyadenylation site.

**Table 1.** Molecular properties of the *AtRecQ*-like genes.

Gene	cDNA/ORF	protein	Exons (no.)	Accession number cDNA/ BAC	Locus number
AtRecQ1	<b>2053</b> /1821	606	11	<b>AJ404470</b> /AC012393	At3g05740
AtRecQ12	<b>2361</b> /2188	705	20	<b>AJ404471</b> /AC007654	At1g31360
AtRecQ13	<b>2533</b> /2142	713	19	<b>AJ404472</b> /ATF8D20	At4g35740
AtRecQ14A	<b>3918</b> /3549	1182	26	<b>AJ404473</b> /ATU95973	At1g10930
AtRecQ14B	<b>3877</b> /3453	1150	26	<b>AJ404474</b> /AC018908	At1g60930
AtRecQ15	<b>2917</b> /2736	911	20	<b>AJ421618</b> /AC079280	At1g27880
AtRecQsim	<b>2988</b> /2577	858	16	<b>AJ404475</b> /AC69556	At5g27680

## RESULTS

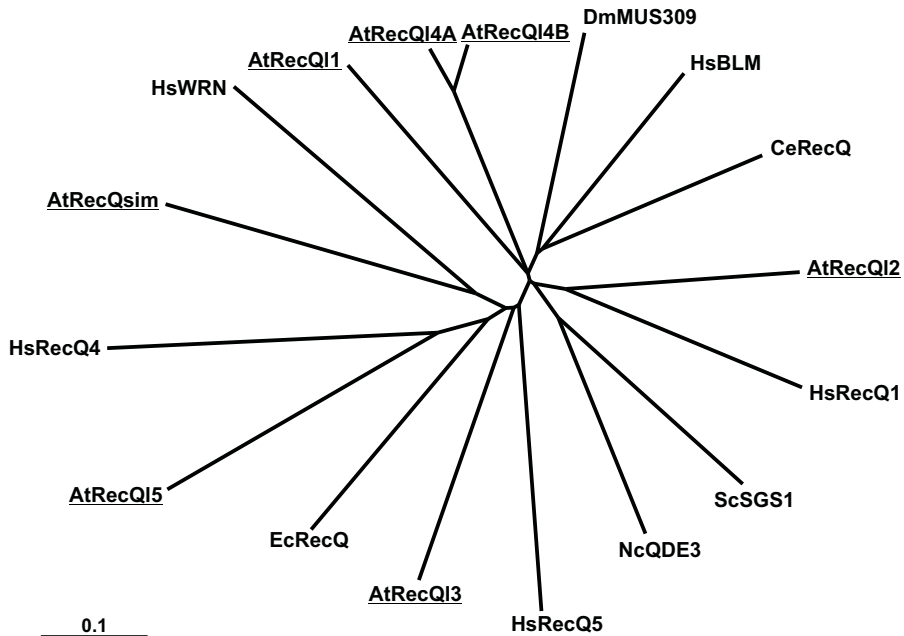
### Identification and analysis of the Arabidopsis *RecQ*-like genes

In order to find the Arabidopsis *RecQ*-like homologues, we used the helicase domain of yeast SGS1 and human WRN proteins to search the Arabidopsis genome. Subsequently, seven putative Arabidopsis homologues were found (Figure 1A and Table 1). The cDNA sequences corresponding to the full-length mRNAs were assembled by making use of RT-PCR and rapid amplification of cDNA ends (RACE). Alignment of the obtained cDNAs with their respective genomic sequences revealed the exon-intron compositions (Figure 1B and Table 1). The results showed that most *AtRecQ*-like genes contain relatively high number of introns, which is rather unusual for Arabidopsis, where 30% of the genes are intronless and the majority has no more than five introns (Delseny and Cooke, 1998). In the region encoding the helicase domain, the number of exons varies between 7 (*AtRecQ13*) and 11 (*AtRecQ11*). The 3<sup>rd</sup> and the 11<sup>th</sup> exons of the *AtRecQ13* gene contain only 8 and 4 base pairs, respectively. Where most *AtRecQ*-like genes exhibit considerable variations in their nucleotide sequence and gene structure, *RecQ14A* and *RecQ14B* are 70% identical at cDNA level and the numbers as well as the positions of their introns are conserved.

Extensive sequence analysis of several independent cDNA clones revealed that the *AtRecQ12* gene encodes for four alternatively processed mRNA species. The most abundant variant was assigned as the full length *AtRecQ12* mRNA (Table1 and Figure 1B). In the second mRNA variant, the second intron is not removed. In the third variant, the acceptor site of the second intron is shifted 49 base pairs downstream. Since these two isoforms carry in frame stop codons before the helicase domain, we concluded that the putative encoded proteins would not be functional helicases. The fourth *AtRecQ12* isoform contains an earlier alternative polyadenylation site and therefore encode for a protein that lacks 130 amino acids from the C-terminus. Similarly, in some *AtRecQsim* cDNAs, the last intron was not removed or an alternative polyadenylation site at position 2368 was found (Figure 1B).

## A

		I		Ia	
AtRecQ11	211	THQACRAS	ER K CFVLM	CGGKSLQ	QLE TLKAS
AtRecQ12	94	REINAT	TC R VVIMAA	CGGKSLQ	QLE MRG
AtRecQ13	41	LEAIGAV	S R CFCLM	CGGKSLQ	QLE LAKP
AtRecQ14A	458	REINAT	S S VFVLM	CGGKSLQ	QLE LIGC
AtRecQ14B	474	REINAT	S C VFVLM	CGGKSLQ	QLE LILCA
AtRecQ15	274	LQAIKMT	CGSST MLVPL	CGGKSLQ	QLE MILPE
AtRecQsim	173	REALSTW	A HK CLVLAA	CGGKSLQ	QLE LL T K
HsRecQ1	56	LETINVT	A C KEVFLVM	CGGKSLQ	QLE LCSD
HsWRN	553	KVHVHSV	EERR NVAVMA	CGGKSLQ	QLE PPVTV
HsBLM	672	LEAINAA	TC E CFVLM	CGGKSLQ	QLE CVSP
HsRecQ4	485	ERAVMT	S IST LVLVL	CGGKSLQ	QLE LIXRRSPCL
HsRecQ5	34	ESATMAV	K NK VFVLM	CGGKSLQ	QLE LLAK
DmBLM	742	LQVINAT	L N CFVLM	CGGKSLQ	QLE ILTE
CeBLM	252	KQCCLST	M H TFVLM	CGGKSLQ	QLE VILPE
ScSGS1	683	LEAVNAT	CG K VFVLM	CGGKSLQ	QLE VVKS
NcQDE3	910	LEAINAT	CG K AFVLM	CGGKSLQ	QLE VVRS
EcRecQ	32	EEIIDTV	SC R CLVLM	CGGKSLQ	QLE LLNG
		II		III	
AtRecQ11	300	CKL	Y APPK	KIAGSSSFLET	RLCLDRKGLLAGEV
AtRecQ12	182	LK	Y APPK	YSKSRKMS	LEKCHNAGRLSLIS
AtRecQ13	129	VR	Y APPK	ELIATKG	FMLKRLKHSRGLNLNLA
AtRecQ14A	546	YK	Y APPK	VAKSDSLR	HLHENLSRGLLARFV
AtRecQ14B	562	YK	Y APPK	VAKSESLR	HLHEILNSRLLARFV
AtRecQ15	355	IK	Y APPK	SPERLLNVEF	LSMFRMSLSVSLVV
AtRecQsim	259	MY	Y APPK	Y CPTTV	RLIKPLQKAKTHGIALFA
HsRecQ1	144	LK	Y APPK	IAKSRKMS	FRLEKAYEARFRTRIA
HsWRN	636	YR	Y APPK	YSCGNMGL	QQDLE ADIGITILIA
HsBLM	760	IK	Y APPK	YCASNLIS	TSTENLYERKLLARFV
HsRecQ4	574	YK	Y APPK	YALVAGG	LPPAAQLPVAFAFAC
HsRecQ5	122	YK	Y APPK	YAPPS	MAASSSFQPTILNSVSRHLISYL
DmBLM	830	YK	Y APPK	YISSSAR	FQDITDLNSNNYSIRFV
CeBLM	340	IK	Y APPK	YKISASGR	LSNVFFDLHRRGLLARFV
ScSGS1	773	LD	Y APPK	YSPMISAS	EGCKRAISRLYADGKLARIV
NcQDE3	1003	EL	Y APPK	YAPMVSKN	QTFVNMKMDLYRRKLIARIV
EcRecQ	118	IR	Y APPK	YRRLMDLN	FLEHLAHNPV
		IV		V	
AtRecQ11	383	KS	Y APPK	YPRAP	VLKMFD
AtRecQ12	265	EM	Y APPK	YH	PKCV
AtRecQ13	211	DS	Y APPK	YRNPL	VLKS
AtRecQ14A	629	QAC	Y APPK	YVNCV	VFRO
AtRecQ14B	645	QAC	Y APPK	YVNCV	VFRO
AtRecQ15	440	SS	Y APPK	YPT	NLIQK
AtRecQsim	345	ES	Y APPK	Y	SKETKIVLT
HsRecQ1	227	KIC	Y APPK	Y	KECP
HsWRN	716	RC	Y APPK	Y	RNPQ
HsBLM	843	TC	Y APPK	Y	KLRLP
HsRecQ4	654	QHA	Y APPK	Y	AEEPDLHGPAVPPTN
HsRecQ5	205	AA	Y APPK	Y	HKKPVAIFKTECF
DmBLM	913	AC	Y APPK	Y	HNKCK
CeBLM	425	DR	Y APPK	Y	KQNSK
ScSGS1	856	HN	Y APPK	Y	PEFV
NcQDE3	1087	HN	Y APPK	Y	ADQC
EcRecQ	196	RI	Y APPK	Y	NDFL
		VI		VII	
AtRecQ11	417	QEL	Y APPK	Y	RD
AtRecQ12	302	AE	Y APPK	Y	RE
AtRecQ13	247	GNL	Y APPK	Y	KSCG
AtRecQ14A	663	DKF	Y APPK	Y	KE
AtRecQ14B	679	DKF	Y APPK	Y	RE
AtRecQ15	473	D	Y APPK	Y	LI
AtRecQsim	440	AK	Y APPK	Y	NSNGK
HsRecQ1	264	VKL	Y APPK	Y	NG
HsWRN	754	KT	Y APPK	Y	SSHWEF
HsBLM	878	E	Y APPK	Y	WRK
HsRecQ4	687	A	Y APPK	Y	IT
HsRecQ5	247	K	Y APPK	Y	GOEADK
DmBLM	948	S	Y APPK	Y	RY
CeBLM	459	V	Y APPK	Y	BEK
ScSGS1	890	C	Y APPK	Y	DA
NcQDE3	1121	A	Y APPK	Y	EL
EcRecQ	226	D	Y APPK	Y	QEL
		VIII		IX	
AtRecQ11	452	KVK	Y APPK	Y	TVY
AtRecQ12	337	G	Y APPK	Y	ISADY
AtRecQ13	280	G	Y APPK	Y	ISSAA
AtRecQ14A	697	G	Y APPK	Y	IKAA
AtRecQ14B	713	G	Y APPK	Y	IKAA
AtRecQ15	510	N	Y APPK	Y	INAK
AtRecQsim	521	G	Y APPK	Y	ILKAA
HsRecQ1	298	G	Y APPK	Y	ILKAA
HsWRN	788	N	Y APPK	Y	ILSCG
HsBLM	912	G	Y APPK	Y	ILKAA
HsRecQ4	738	T	Y APPK	Y	ILKAA
HsRecQ5	283	G	Y APPK	Y	ILKAA
DmBLM	982	G	Y APPK	Y	ILKAA
CeBLM	493	G	Y APPK	Y	ILKAA
ScSGS1	924	G	Y APPK	Y	ILKAA
NcQDE3	1156	R	Y APPK	Y	ILKAA
EcRecQ	262	G	Y APPK	Y	ILKAA

**B**

**Figure 2.** Evolution of *AtRecQ*-like genes. (A) Aligned amino acid residues of the helicase domain of 17 RecQ-like proteins. The alignment was made using the Clustal W method, as described in Materials and methods. Roman numbers indicate the seven motifs of the helicase domain. (B) A Phylogram of RecQ proteins indicated in figure 2A generated by using the Phylip method. Branch distances correspond to sequence divergence and the length of each pair of branches represents the distance between sequence pairs. At, *Arabidopsis thaliana*; Hs, *Homo sapiens*; Sc, *Saccharomyces cerevisiae*; Nc, *Neurospora crassa*; Ce, *Caenorhabditis elegans*; Dm, *Drosophila melanogaster*; Ec, *Escherichia coli*.

### Evolution of the Arabidopsis *RecQ* family members

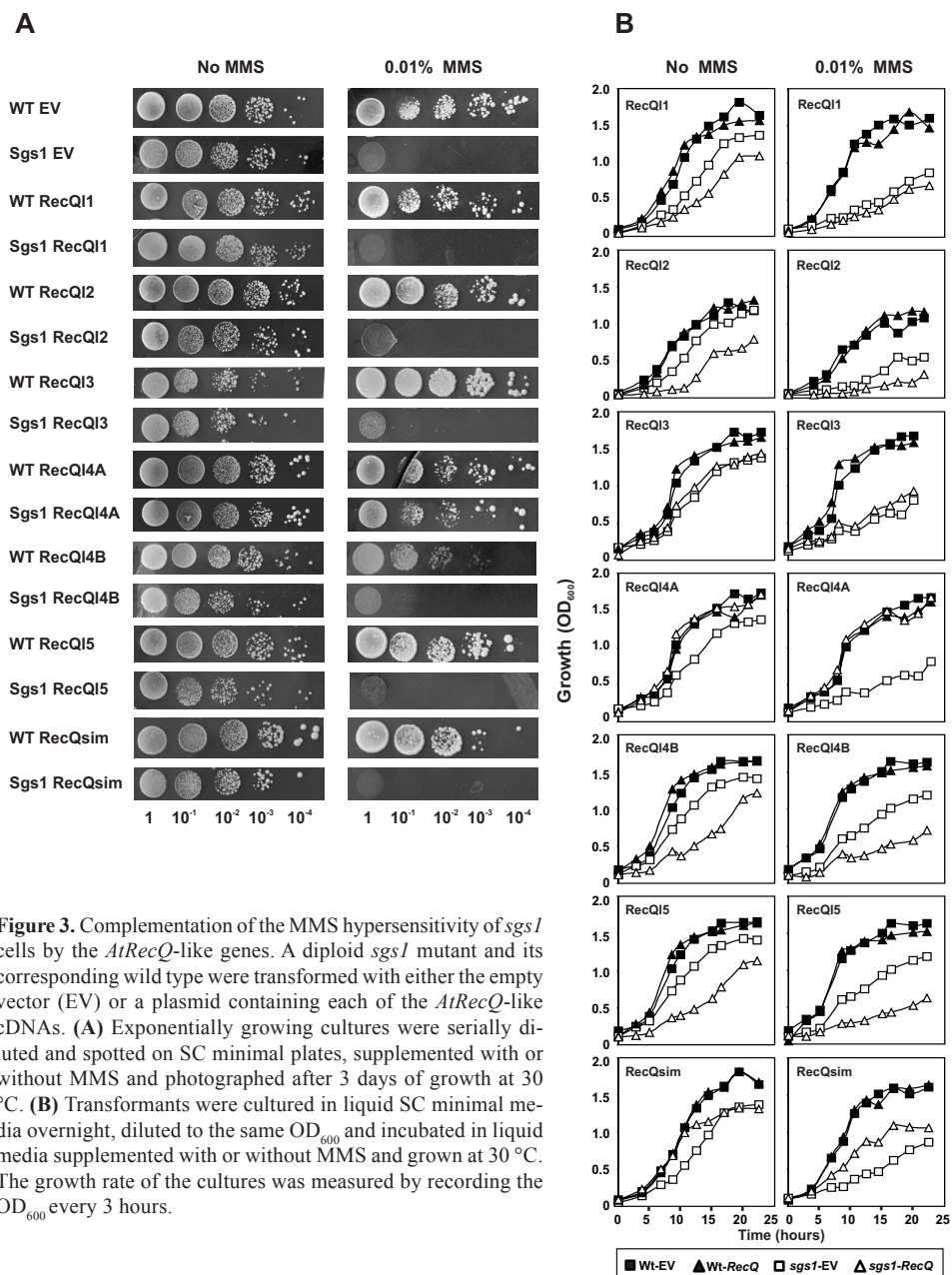
The evolutionary relationship of the Arabidopsis *RecQ* homologues was further analysed at the protein level. The deduced amino acid sequences of the helicase domain of *AtRecQ*-like proteins were aligned with that of the five known human *RecQ* homologues, together with representative members of different organisms, as described in Materials and methods. The result presented in Figure 2A revealed that the seven conserved motifs of the helicase domain of all *AtRecQ*-like proteins show a clear homology to that of the other *RecQ* family members. Among the Arabidopsis *RecQ* proteins, the identity at the amino acid level within the helicase domain is between 30% (*AtRecQ15* and *AtRecQsim*) and 94% (*AtRecQ14A* and *AtRecQ14B*). Interestingly, the helicase domain in Arabidopsis *RecQsim* and human *RecQ4* proteins contains an insert of about 100 and 13 amino acids, respectively (Figure 2A).



The alignment was employed to establish a phylogenetic tree analysis. The resulting tree is shown in Figure 2B and indicates that most Arabidopsis RecQ proteins are more close to non-plant RecQ members than to each other. Arabidopsis RecQ12 and RecQ13 are more close to human RecQ1 and RecQ5, respectively. Arabidopsis RecQsim and RecQ15 are more similar to human WRN and RecQ4, respectively. AtRecQ11 encompass a separate subgroup and AtRecQ14A and 4B are very closely related. The data suggest that these two homologues have evolved from a recent duplication in plants while the other plant *RecQ* homologues have evolved before the divergence of plants and animals.

### **Some Arabidopsis *RecQ*-like genes can enhance the slow growth of *sgs1* mutant in the absence or presence of genotoxic stress**

Under standard growth conditions, budding yeast *sgs1* cells are viable and if grown on solid media their plating efficiency is similar to that of the wild type. In liquid culture, however, *sgs1* cells exhibit a moderate increase in the doubling time as compared to wild type (Gangloff et al., 2000; Branzei et al., 2002). When HU and MMS is added to the culture, *sgs1* cells show more sensitivity than wild type (Yamagata et al., 1998; Frei and Gasser, 2000). The HU hypersensitivity can be suppressed by ectopic expression of the human *BLM* gene (Yamagata et al., 1998). We tested whether any of the Arabidopsis *RecQ* family members could suppress the MMS hypersensitivity of the *sgs1* mutant. The *AtRecQ* cDNA fragments containing the complete ORF were sub cloned into a yeast vector under the control of the *GAL1* promoter and a diploid *sgs1* strain along with its respective wild type were transformed with either the resulting construct or the empty vector. Subsequently, the MMS hypersensitivity of the transformants was monitored by the spot assay on solid media, as described by Ui et al. (2001). As shown in Figure 3A, in the absence of MMS all clones were able to grow on plates lacking uracil with a similar apparent plating efficiency. In the presence of 0.01% MMS, none of the transformants with the wild type background showed hypersensitivity to MMS. Thus, we concluded that the ectopically expressed *AtRecQ* genes did not interfere with the wild type *SGS1* in conferring MMS resistance. In contrast, most transformants with *sgs1* background showed hypersensitivity to MMS. Consistent with our results presented in chapter 4, this hypersensitivity was restored by Arabidopsis *RecQ14A*, suggesting that *RecQ14A* can fulfil the role of *SGS1* in yeast. The roles that Arabidopsis *RecQ* genes may play in yeast were further analysed by monitoring the growth curves of the transformants in SC minimal liquid media, lacking uracil, and supplemented with or without MMS. The results that are shown in Figure 3B generally confirmed the spot assay data. The growth curve analysis, in addition, showed that even in the absence of MMS, *sgs1* cells transformed with the empty vector grew slightly slower than the corresponding wild type, consistent with previous observations (Gangloff et al., 2000; Branzei et al., 2002). Interestingly, this slow growth was fully restored by *RecQ14A* and was partially suppressed by *RecQsim*. Indeed, *sgs1* cells transformed with *AtRecQsim* grew better than *sgs1* cells in the first 20 hours of growth but the difference disappeared afterwards. In addition, *RecQ13* did not have an obvious effect, while *RecQ11*, *RecQ12*, *RecQ14B* and *RecQ15* enhanced the slow growth phenotype of the *sgs1* cells (Figure 3B). Comparable results were observed when 0.01% MMS was added to the medium. Thus, we concluded that this effect is independent of DNA damage induced by MMS. Collectively, the data



indicate that the Arabidopsis *RecQ* genes may play different roles in *sgs1* cells and suggest that *RecQ14A* is the Arabidopsis orthologue of *SGS1*.

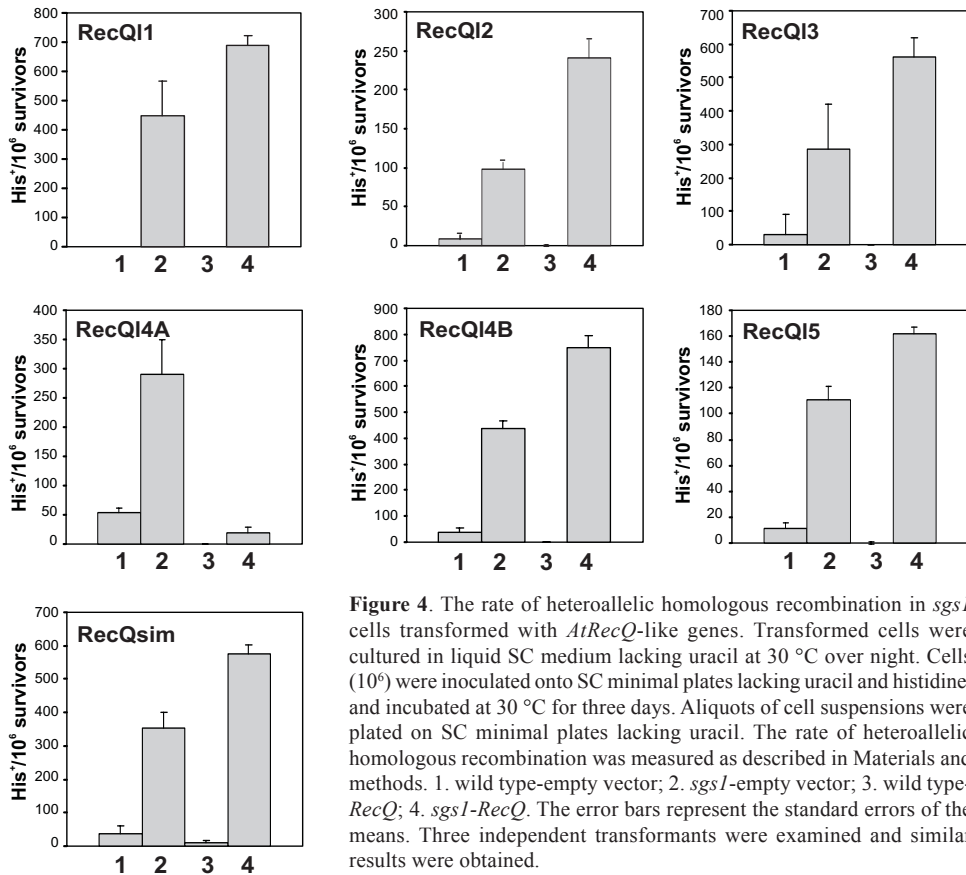
### **Most *AtRecQ*-like genes increase the rate of homologous recombination in *sgs1* cells**

In addition to MMS hypersensitivity, *sgs1* cells show genetic hyper-recombination phenotypes including enhanced interchromosomal homologous recombination (Yamagata et al., 1998). Therefore, the rate of homologous recombination in the strains indicated in figure 3 was studied. The yeast strains used in this study carry a *his1-1/his1-7* mutation, which allows detection of interchromosomal heteroallelic homologous recombination by the appearance of histidine (His) prototrophy (Ui et al., 2001). Exponentially growing cultures were diluted and grown on plates supplemented with or without histidine. The rate of heteroallelic interchromosomal homologous recombination was measured by calculating the restoration rate of histidine prototrophy. As expected, the data presented in figure 4 show that the *sgs1* strain exhibited accelerated interchromosomal HR. The rate of interchromosomal HR was not dramatically changed in the wild type cells transformed with either the empty vector or the *RecQ* genes. Consistent with data presented in chapter 4, *RecQ14A* restored the enhanced recombination rate of *sgs1* cells but all other Arabidopsis *RecQ*-like genes did the opposite. The data support the idea that most *AtRecQ*-like genes perform functions that differ from the function of *SGS1*.

## **DISCUSSION**

Association of mutations in human *RecQ* genes with three distinct heritable diseases that manifest in genome instability and cancer predisposition brought wide attention to elucidate the precise functions of *RecQ* family members in various model organisms. Here, we report that the Arabidopsis genome contains seven *RecQ*-like genes, which is in agreement with the results published during the course of this work (Hartung et al., 2000; accession number AJ421618). The *RecQ* cDNAs isolated in our lab were generally larger than what has been reported, however, those differences would not result in protein-length or sequence alterations.

We found that some Arabidopsis *RecQ*-like genes encode for alternatively processed mRNA molecules. This, together with ubiquitous expression of *AtRecQ*-like genes (Hartung et al., 2000; Bagherieh-Najjar et al., 2003), implies that alternative mRNA processing might play a role in posttranscriptional regulation of some Arabidopsis *RecQ*-like genes. Similarly, the human and *Drosophila RecQ5* genes encode for alternatively processed mRNA species (Sekelsky et al., 1999; Ozsoy et al., 2001). In man, these alternatively processed transcripts exhibit different sub cellular localisation and perhaps different functions (Shimamoto et al., 2000). It has been suggested that alternative transcription initiation and alternative splicing is less widely used in plants than in animals (Arabidopsis Genome initiative 2000). However, increasing evidences indicate that many plant genes encode for alternative transcripts, suggesting that alternative mRNA processing might have an under appreciated role in expanding proteom diversity in plants (Lorkovic et al., 2000). Nevertheless, future



**Figure 4.** The rate of heteroallelic homologous recombination in *sgs1* cells transformed with *AtRecQ*-like genes. Transformed cells were cultured in liquid SC medium lacking uracil at 30 °C over night. Cells (10<sup>6</sup>) were inoculated onto SC minimal plates lacking uracil and histidine, and incubated at 30 °C for three days. Aliquots of cell suspensions were plated on SC minimal plates lacking uracil. The rate of heteroallelic homologous recombination was measured as described in Materials and methods. 1. wild type-empty vector; 2. *sgs1*-empty vector; 3. wild type-*RecQ*; 4. *sgs1-RecQ*. The error bars represent the standard errors of the means. Three independent transformants were examined and similar results were obtained.

works are needed to clarify the possible role of alternative mRNA processing in functional regulation of Arabidopsis *RecQ* genes.

*RecQ*-like genes occur in three domains of life *Archaea*, *Bacteria*, and *Eucarya*, but not in all species. Thirteen out of 15 sequenced archaeal genomes and 20 out of 48 sequenced bacterial genomes are devoid of a gene encoding for a *RecQ* helicase (Nakayama, 2002). However, in *Eucarya*, all studied organisms contain at least one *RecQ* gene. Most Arabidopsis *RecQ*-like genes are heterogeneous at the DNA nucleotide sequence level, the compositions of intron-exon boundaries, N- and C-terminal domains, and according to a phylogenetic tree analysis, in most cases, are more close to non-plant *RecQ* proteins than to each other. These observations may be taken to suggest that multiple *RecQ* genes were present in the last common ancestor of plants and animals. Current knowledge based on Kingdom-level phylogenetic analyses implies that the evolutionary divergence between plants and animals is more than that between animals and fungi (Balduaf et al., 2000). Thus, considering the

existence of a single *RecQ* gene in yeast and multicopy *RecQ* genes in plants and animals, it seems likely that evolution had both positive and negative pressures on the number of *RecQ* genes present in the genome of different organisms. Accordingly, in the evolution of animals and plants the number of *RecQ* genes generally increased in parallel with an increase in the complexity of the genomes but in lower organisms they were removed. In agreement, Nakayama (2002) proposed that *RecQ* helicases can be a luxury for the cells and they have been lost in the process of reductive evolution of cells with small genomes. Nevertheless, the presence of distantly related multiple *RecQ* genes in one organism implies that they may have distinct functions.

Functional specification of the Arabidopsis *RecQ*-like genes was further supported by our data obtained from their ectopic expression in the budding yeast. Based on these data, *AtRecQ*-like genes are biologically active in yeast and might be divided in four functional classes. Class 1, *AtRecQ13*, did not have obvious effects on growth rate and MMS hypersensitivity of *sgs1* cells but increased the rate of HR. Class 2, *AtRecQ11*, *AtRecQ12*, *AtRecQ14B*, and *AtRecQ15*, decreased the growth rate of *sgs1* cells and enhanced the rate of HR. Several nonexclusive possible functions for this group of Arabidopsis *RecQ* genes can be proposed. First, they could exhibit a function similar to *SGS1* but their overexpression somehow caused the reverse. Consistently, overexpression of *SGS1* had strong dominant-negative effects in the presence of MMS (Cobb et al., 2003). This scenario, however, seems unlikely because in the same conditions the ectopic expression of *AtRecQ*-like genes in wild type cells did not have adverse effects. Secondly, they might function in aspects of DNA metabolism, as enhancers of HR or suppressors of NHEJ, which is activated only in the absence of a functional *SGS1*. In support of this notion, *E. coli* RecQ was shown to be a suppressor of NHEJ and only in a *recBC sbcB* background was essential for HR (Hanada et al., 1997). Alternatively, Topoisomerase III (Top3), a well-characterised partner of *SGS1* might play a role in the functionality of this group of *AtRecQ*-like genes. *TOP3* deficiency in budding yeast leads to a dramatic slow growth, which can be suppressed by inactivation of *SGS1* (Gangloff et al., 1994). However, the rate of HR, which in *top3* mutant is less than that of wild type, is dramatically increased in *sgs1 top3* double mutant to a level more than that in *sgs1* single mutant (Myung et al., 2001). If the product of this class of *AtRecQ* genes could interact with Top3, and somehow compromise its function, the rate of HR would be upregulated only in *sgs1* cells and not in the wild type, consistent with our observation.

The third class of the Arabidopsis *RecQ*-like genes, which only contains *RecQ14A*, was able to almost completely suppress the hyper-recombination, the MMS hypersensitivity, and the slow growth rate of *sgs1* mutant, suggesting that *RecQ14A* is the plant orthologue of *SGS1*. It was however surprising that *AtRecQ14B* with more than 84% overall similarity (67% identity) at amino acid level with *AtRecQ14A* behaved differently. This suggests that after gene duplication they acquired different functions. Functional diversion after gene duplications is considered as a common evolutionary mechanism in higher organisms (Arabidopsis genome initiative, 2000). Three observations support the idea that these two genes perform different roles. (i) The expression profile of *RecQ14A* and *RecQ14B* is not similar in Arabidopsis (Bagherieh-Najjar et al., 2003). (ii) When expressed in yeast cells, *RecQ14A* was able to compensate for the absence of *SGS1* but *RecQ14B* could not (this report and Chapter 4). (iii) Arabidopsis *recQ14A* mutant plants show altered responses to genotoxic stress and increased rate of homologous recombination (see Chapter 4), indicating that *AtRecQ14B* cannot fulfil the role that *AtRecQ14A* plays in plant metabolism. Nevertheless,

phenotypic analysis of *recQ14B* single and *recQ14A/recQ14B* double mutant plants will clarify the role of these two genes more precisely.

The only member of the fourth functional class of *AtRecQ* genes, *RecQsim*, was able to partially suppress the MMS hypersensitivity of the *sgs1* cells in liquid culture but enhanced the increased rate of HR. Ectopic expression of *AtRecQsim* in a haploid *sgs1* strain suppressed the MMS hypersensitivity of the mutant cells, as well (Bagherieh-Najjar et al., 2003). However, the plating efficiency of diploid *sgs1 RecQsim* cells in the presence of MMS was less than that of haploid cells. This might be related to various levels of MMS hypersensitivity of different strains used in these studies. In agreement, Bachrati and Hickson (2003) indicated that there are significant differences in the MMS hypersensitivity of different *sgs1* strains.

We showed that the role of *RecQsim* in suppression of MMS hypersensitivity of *sgs1* cells opposed its role in enhancement of HR. This observation suggests that the increased rate of HR and the slow growth phenotype of *sgs1* cells can be independent. Consistently, ectopic expression of human WRN suppressed the rate of recombination of *sgs1* cells without any effect on the slow growth of the mutant cells in presence of HU (Yamagata et al., 1998). Recent data suggest that in the presence of hydroxyurea, SGS1 helicase resolves aberrantly paired structures at stalled DNA replication forks to maintain single-stranded DNA that allows stabilization of DNA polymerase at the stalled forks (Cobb et al., 2003). Inactivation of *SGS1*, therefore, might lead to both the activation of DNA damage checkpoints associated with slow growth and to increased rate of recombination in aberrantly structured DNA molecules formed at stalled replication forks. Analogous pathways in plants may occur which involve some of Arabidopsis *RecQ*-like genes.

## MATERIALS AND METHODS

### Plant materials and growth conditions

Seeds of *Arabidopsis thaliana* accession Columbia (Col-0) were sown directly onto an organic-rich soil (Tulip profi No.4, Bogro B.V., The Netherlands). The trays were kept at 4 °C in darkness for 4 days and transferred to a growth chamber at 21 °C and 65% relative humidity with 16 h light ( $\sim 60 \mu\text{mol m}^{-2} \text{s}^{-1}$ )/8 h dark cycles.

### RNA isolation, Northern analysis, reverse transcription PCR and RACE

Total RNA was isolated from the aerial parts of the plant as described by Sambrook and Russell (2001). The reverse transcription (RT) reactions were performed in a total volume of 20  $\mu\text{l}$ , with 2  $\mu\text{g}$  of total RNA and 0.2  $\mu\text{M}$  oligo-dT primers, using the Omniscript RT kit (Qiagen, Germany), following the instructions of the supplier. The PCR reactions (total volume of 50  $\mu\text{l}$ ) contained 0.2 mM of each dNTPs, 0.2  $\mu\text{M}$  of each primer, 4  $\mu\text{l}$  of the 4 $\times$  diluted RT reactions, 1 $\times$ Taq polymerase buffer, and 1 unit of Taq DNA polymerase (Roche Diagnostics). The cDNA ends were obtained, using the 5'/3' RACE kit (Roche) following the instructions of the supplier. The nucleotide sequences of the gene specific primers are available upon request. The amplified fragments were cloned in the pGEMT-EASY vector (Promega), for sequencing.

### Computer analysis

The BLAST searches were performed, using the TAIR and NCBI BLAST programs at <http://Arabidopsis.org/> and <http://ncbi.gov/>, respectively. The alignment and the respective phylogenetic tree were performed at <http://www.ebi.ac.uk/clustalw/>, using the ClustalW method, scoring matrix of Blosum, gap penalty of 10, and set to generate Phylip phylogenetic tree.

### Yeast complementation

The diploid wild-type MR101 (*MATa/MAT $\alpha$  ura3-52/ura3-52 leu2-3,112/leu2-3,112 trp-298/trp-298 his1-1/his1-7*) and *sgs1* (ds1, MR101 *sgs1::AUR1/sgs1::AUR1*) strains used in this study were generous gifts from T. Enomoto (Ui et al., 2001). Amplification and cloning of Arabidopsis *RecQ* cDNAs into the pYES2/CT (Invitrogen) is described in Chapter 4. All subsequent experiments with the transformants were conducted on media lacking uracil (Ura). The MMS spot assay and measurement of the frequency of interchromosomal recombination between heteroalleles, *his1-1/his1-7*, was performed as described in Chapter 4.

### ACKNOWLEDGEMENTS

We are grateful to T. Enomoto for providing the yeast strains. We thank W. Meijer, A. Mulder, K. Kostyn and C. Angheluta for excellent technical assistance. This work was supported by a grant from the Ministry of Science, Research and Technology of the Islamic Republic of Iran to M.B.B.-N.